次元最小の十分統計量の構成 (Invariant Statistical Problemsと Sufficiency)

清水, 良一

数理解析研究所講究録
46 巻 1968 年 15-25

次元最小の十分統計量の構成

統数研　　清水良一

E.W. Barankin, "Sufficient statistics in the case of non-constant carrier," Sankhyā, ser. A., vol. 28., 1966, pp 101－122 を紹介する.

E^n の開集合 Ω の上の確率密度の族,

$$\mathcal{P} = \{ p(x, \theta) \mid \theta \in \Theta \}$$

が与えられているとする. 母数空間 Θ も E^ν の開集合で, すべての $\theta \in \Theta$ について,

$$\Omega = \Omega_\theta = \{ x \mid p(x, \theta) > 0 \}$$

であること, すなわち, $x \in \Omega$, $\theta \in \Theta$ のとき, $p(x, \theta) > 0$ であることを仮定する. さらに, $p(x, \theta)$ が x および θ についてなめらかであるとして, $\mathcal{P}$ にたいする十分統計量 $T(x) = (T_1(x), \cdots, T_r(x))$ で, $T_i(x)$ がなめらかな実数値関数であるもののうち, r が最小のものを構成するという問題は, 前論文, E.W. Barankin and M. Katz, Jr., "Sufficient statistics of minimal dimension," Sankhyā, vol. 21, 1959, pp 217－246,

で解決された.　　この場合は, $\log p(x, \theta)$ が定義され, $\operatorname{grad}_\theta \log p(x, \theta)$ が最小十分統計量になるので, 問題は $\operatorname{grad}_\theta \log p(x, \theta)$, $\theta \in \Theta$ で張られる linear space の基を求めることに帰着した.

この論文では, $\Omega_\theta = \{x \mid p(x, \theta) > 0\}$ が θ に依存しうる場合に同じ問題を扱う.　以下, つぎの仮定を設ける.

(I)　$\Gamma = \bigcup_{\theta \in \Theta} (\Omega_\theta \times \{\theta\})$ が $E^{n+\nu}$ の開集合である.

(II)　$p(x, \theta)$ が Γ で連続である.

(III)　Γ の境界のルベーグ測度は 0.

(IV)　Γ の境界の θ-section はルベーグ測度 0.

(V)　$x \in \Omega$ にたいして, 有限の値 $b(x)$ があって, すべての $\theta \in \Theta$ について, $p(x, \theta) \leqq b(x)$.

さて, $m = 1, 2, \cdots$ にたいして, $\varphi_m(\theta, \tau)$ は $\Theta \times \Theta$ で定義され, 正の値をとる可測関数で, $\theta \in \Theta$ にたいして, $\varphi_m(\theta, \cdot)$ が Θ 上の確率密度であるとする.

$$(1)\quad p_m(x, \theta) = \int_\Theta p(x, \tau)\varphi_m(\theta, \tau)\, d\tau$$

とおくと，　$\theta\in\Theta$ にたいして，$p_m(\cdot,\theta)$ が Ω 上の確率密度であり，しかも，すべての $x\in\Omega$，$\theta\in\Theta$ にたいして，

$p_m(x,\theta)>0$　である．　　したがって，$p(\cdot,\theta)$ がなめらかならば，（θ についての微分可能性を仮定しなくても），$\varphi_m(\theta,\tau)$ に適当な条件を課すことによって，$p_m(x,\theta)$ がなめらかとなり，族 $\mathcal{P}_m=\{p_m(x,\theta)\mid\theta\in\Theta\}$ に前論文の結果を適用することができる．　　こゝでは，系列 $\varphi_m(\cdot,\cdot)$ として，つぎのようなものを考える．

E^ν において，θ を中心とする半径 α の球の内部を $S_{\theta,\alpha}$ とする．　このとき，任意の $\alpha>0$ にたいして，

$$(2)\qquad \lim_{m\to\infty}\int_{\Theta-S_{\theta,\alpha}}\varphi_m(\theta,\tau)\,d\tau=0 \qquad \theta\in\Theta$$

これは，確率密度 $\varphi_m(\theta,\cdot)$ をもつ分布の列が，θ に確率 1 を与える分布に収束することを意味する．

この条件のもとで，

$$(3)\qquad \lim_{m\to\infty}p_m(x,\theta)=p(x,\theta) \qquad x\in\Omega,\ \theta\in\Theta$$

が成り立つ．　　実際，

$$|p_m(x,\theta)-p(x,\theta)| \leqq \int_{\Theta} \varphi_m(\theta,\tau)|p(x,\tau)-p(x,\theta)|\,d\tau$$

であるが，$p(x,\cdot)$ の連続性から，任意の $\varepsilon>0$ にたいして，$\alpha>0$ を十分小さくとるとき，$\tau\in S_{\theta\alpha}$ なら，$|p(x,\tau)-p(x,\theta)| \leqq \varepsilon$ となるから，積分領域を $\Theta-S_{\theta,\alpha}$ と，$S_{\theta,\alpha}$ とに分け，さらに（V）を考慮すると，

$$|p_m(x,\theta)-p(x,\theta)| \leqq 2b(x)\int_{\Theta-S_{\theta,\alpha}} \varphi_m(\theta,\tau)\,d\tau+\varepsilon.$$

よって（2）から $\varlimsup_{m\to\infty}|p_m(x,\theta)-p(x,\theta)| \leqq \varepsilon$．$\varepsilon$ が任意だから（3）が導かれる．

定理 1．　$\varphi_m(\theta,\tau)$ が上記の条件を満たすとする．統計量 $T(x)$ が $\mathcal{P}$ にたいして十分であるための必要十分条件は，$T(x)$ が $\mathcal{P}_m$，$m=1,2,\cdots$ にたいして十分なことである．

証明．　$T(x)$ が $\mathcal{P}$ にたいして十分であるとする．分解定理によって，

$$p(x,\theta)=f(T(x),\theta)\cdot g(x)$$

と書けるが，この関係は $x\in\Omega-M$ の各々について，ほとんどすべての $\theta\in\Theta$ にたいして成り立つこと

が示される．　（こゝおよび以下において，M は，ルベーグ測度 0 の，Ω の適当な部分集合を表わす．）

よって，

$$p_m(x,\theta)=\int_{\Theta} f(T(x),\theta)\,g(x)\cdot\varphi_m(\theta,\tau)\,d\tau.$$

$$=f_m(T(x),\theta)\cdot g(x). \qquad \text{a.e. } \theta \text{ for each } x\in\Omega-M.$$

たゞし，　$f_m(y,\theta)=\int_{\Theta} f(y,\theta)\,\varphi_m(\theta,\tau)\,d\tau.$

逆に　$T(x)$ が $\mathcal{P}_m$，$m=1,2,\cdots$　にたいして十分であるとする．　分解定理によって，

(4)　　$$p_m^{\circ}(x,\theta)=f_m^{\circ}(T(x),\theta)\cdot g_m^{\circ}(x).$$

とかける．

$$\Omega=\bigcup_{\theta\in\Theta}\Omega_\theta$$

であるが，Ω_θ がすべて開集合だから，$\theta^{(k)}\in\Theta$, $k=1,2,\cdots$ が存在して，

$$\Omega=\bigcup_{k=1}^{\infty}\Omega_{\theta^{(k)}}$$

と書ける．

$$a_n>0,\qquad \sum_n a_n=1.$$

として，

(5) $$g_m(x) = \sum_k a_k p_m(x, \theta^{(k)}) \quad , \qquad x \in \Omega$$

とおくと,

$$g_m(x) > 0, \qquad m = 1, 2, \cdots, \quad x \in \Omega$$

であり，また,

$$\int_\Omega g_m(x)dx = \sum a_k \int_\Omega p_m(x, \theta^{(k)})dx = 1.$$

から, $g_m(x) < \infty$ である. (4) を (5) に代入して,

(6) $$g_m(x) = g_m^\circ(x) \cdot \sum_k a_k \cdot f_m^\circ(T(x), \theta^{(k)}).$$

となるから,

$$f_m(y, \theta) = f_m^\circ(y, \theta) / \sum_k a_k f_m^\circ(y, \theta^{(k)})$$

とおくと, (4) は,

(7) $$p_m(x, \theta) = f_m(T(x), \theta) \cdot g_m(x)$$

と書ける. 定理の仮定から,

(8) $$\lim_{m \to \infty} p_m(x, \theta) = p(x, \theta) \quad , \qquad x \in \Omega, \ \theta \in \Theta$$

となるから, (5) と bounded convergence theorem から

$$\lim_{m \to \infty} g_m(x) = \lim_m \sum_k a_k p_m(x, \theta^{(k)}) = \sum_k a_k p(x, \theta^{(k)})$$

$$\equiv g(x) \quad > 0, \qquad x \in \Omega$$

が存在する．　よって，$x \in \Omega - M$ のとき，a.e. θ について，

$$p(x,\theta)/g(x) = \lim_{m\to\infty} p_m(x,\theta)/g_m(x) = \lim_{m\to\infty} f_m(T(x),\theta)$$

が $T(x)$ と θ の関数 $f(T(x),\theta)$ に等しい．

$$p(x,\theta) = f(T(x),\theta)\cdot g(x)$$

この関係が，実は，すべての $\theta \in \Theta$ について，a.e. x で成り立つことが示され，$T(x)$ が $\mathcal{P}$ にたいして十分であることが分る．　q.e.d.

例．　$\Theta = (0, \infty)$，$x = (x_1, \cdots, x_n)$

$$p(x,\theta) = \begin{cases} 1/\theta^n & \text{if } 0 < x_i < \theta, \ i = 1, \cdots, n \\ 0 & \text{otherwise} \end{cases}$$

とする．　$\Omega = [(0,\infty)\times\cdots\times(0,\infty)] - E$　としてよい．　ただし，$E = \bigcup_{i\neq j}\{(x_1,\cdots,x_n) \mid x_i = x_j\}$ とする．
$\Omega_\theta = [(0,\theta)\times\cdots\times(0,\theta)] - E$　である．

ここでは，仮定（I）―（IV）が満たされることは明らか。（V）は，$\theta \leqq \max x_i$ なら $p(x,\theta) = 0$，$\theta > \max x_i$ なら $p(x,\theta) = 1/\theta^n$ であるから，$\theta \in \Theta$，$x \in \Omega$ にたいして，

$$p(x,\theta) \leqq (\max x_i)^{-n}$$

である。　いまの場合，

$$p_m(x,\theta)=\int_{\min x_i}^{\infty}\frac{1}{\tau^n}\varphi_m(\theta,\tau)\,d\tau$$

となるが，これが θ についてなめらかになるように φ_m を選べば，前論文の結果が適用される。それによれば，Ω の点はすべて regular となり，各点での十分統計量の最小次元は 1 であることが分る。さらに分解定理から，$T(x)=\max x_i$ が P_m, $m=1,2,\cdots$ の，したがって P の十分統計量であることが分る。

系列 $\varphi_m(\theta,\tau)$ の代りに，たゞ１つの $\varphi(\theta,\tau)$ を用いて議論することも可能である。$\varphi(\theta,\tau)$ は，$\varphi_m(\theta,\tau)$ と同様の性質をもつ他，"Θ_0 がルベッグ測度 0 の，Θ の部分集合のとき，

$$\{\varphi(\theta,\cdot)\mid\theta\in\Theta-\Theta_0\}$$

が $L_1(\Theta)$（$=\Theta$ 上の絶対可積分な実数値関数のつくる空間）で完備である。" という条件を満たすとする。

$$\mathcal{Q}=\left\{q(x,\theta)=\int_{\Theta}\varphi(\theta,\tau)p(x,\tau)\,d\tau\;\middle|\;\theta\in\Theta\right\}$$

とおく。　$q(x,\theta)>0$,　$x\in\Omega$,　$\theta\in\Theta$ である

定理 2.　φ が上記の条件を満たすとする．統計量 $T(x)$ が $\mathcal{P}$ にたいして十分であるための必要十分条件は，$T(x)$ が $\mathcal{Q}$ にたいして十分なことである．

証明　必要であることの証明は定理1と同じである．十分であること：　$T(x)$ が $\mathcal{Q}$ にたいして，十分であるとする．分解定理から，

$$(9) \qquad q(x,\theta) = f(T(x),\theta)\, g(x), \qquad \text{a.e. } \theta \text{ for each } x \in \Omega - M$$

と書けるが，前と同様，$0 < g(x) < \infty$ としてよい．
$q(x,\theta)$ の定義と (9) から

$$(10) \qquad \int_{\Theta} \varphi(\theta,\tau)\, p(x,\tau)\, d\tau = f(T(x),\theta)\, g(x).$$

$$\text{a.e. } \theta \quad \text{for each } x \in \Omega - M.$$

いま，$x, y \in \Omega - M$，$T(x) = T(y)$ とすると，
(10) から

$$\int_{\Theta} \varphi(\theta,\tau)\left[g(x)\,p(y,\tau) - g(y)\,p(x,\tau)\right] d\tau = 0 \qquad \text{a.e. } \theta$$

よって，

$$g(x)\,p(y,\theta) - g(y)\,p(x,\theta) = 0. \qquad \text{a.e. } \theta$$

すなわち，

$$p(x,\theta)/g(x) = p(y,\theta)/g(y)$$

これは $p(x,\theta)/g(x)$ が $T(x)$ と θ との関数，$f(T(x),\theta)$ に等しいことを示す．

$$p(x,\theta) = f(T(x),\theta)\,g(x).$$

この関係が，実はすべての $\theta \in \Theta$ について，a.e x で成り立つことがいえて，$T(x)$ が $\mathcal{P}$ にたいする十分統計量であることが分る．　　　q.e.d.

Remark.　$\varphi(\theta,\tau)$ が上記の条件すべてを満足するためのひとつの十分条件は，

$$\varphi(\theta,\tau) = \prod_{i=1}^{\nu} h(\theta_i - \tau_i)$$

と書けることである．ただし，$\theta = (\theta_1,\cdots,\theta_\nu)$，$\tau = (\tau_1,\cdots,\tau_\nu)$ で h は $\mathcal{R}$ 上の連続な確率密度で，いたるところ正，しかもその特性関数が 0 にならないものとする．

例　$x = (x_1,\cdots,x_n)$，$\Theta = \mathcal{R}$，

$$p(x,\theta) = \begin{cases} \left(\frac{2}{\sqrt{2\pi}}\right)^n e^{-\frac{1}{2}\sum_1^n (x_i-\theta)^2} & x_i > 0 \\ & i=1,\cdots,n \end{cases}$$

$$\Big\{\ 0 \qquad\qquad\qquad \text{その他.}$$

とする.

正規分布 $N(0,1)$ の特性関数 $e^{-t^2/2}$ が 0 にならないことから,

$$\varphi(\theta,\tau)=\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}(\theta-\tau)^2}$$

が条件を満たすことが分る.

$$q(x,\theta)=\frac{2^n}{(2\pi)^{\frac{n+1}{2}}}\,\exp\left\{-\frac{1}{2}\left[\sum_i(x_i-\bar{x})^2+\frac{n}{n+1}(\theta-\bar{x})^2\right]\right\}\times$$

$$\times\int_{-\infty}^{\min x_i-\frac{\theta+n\bar{x}}{n+1}} e^{-\frac{n+1}{2}y^2}\,dy.$$

に前論文の結果が使えて,$(\min x_i,\ \bar{x})$ が次元最小の十分統計量であることが分る.